

1 CCCCCTGGAC AGCGGCCCTC TCCTCCAGGC CCGTGGGGCT GGCCCTGCAC CGCGAGCTT CCGGGATGA GGGCCCCCGG TGTGTCACC CGCGCGCCCC  
GGGGACCTG TCGCGGGAG AGGAGGTCCG GGCACCCCGA CCGGACGTG GGGCTCGAA GGGCCCTACT CCCGGGGGCC ACACCACTGG GCCCGGGGG  
101 CAGGTCGCTG AGGACCCCGG GCCAGGCGCG GAGATGGGGG TGCACGGTGA GTACTCGCG GCTGGCGCT CCCGCCCCCG CCGGTCCCTG TTTGAGCGGG  
GTCCAGCGAC TCCCTGGGC CCGTCCGCGC CTCACCCCC ACGTGCCACT CATGAGCGC CGACCCGGA GGGCGGGCGG GCCCAGGGAC AAACCTCGGCC  
1 M G V H  
^Exon 1  
^Met  
201 GATTTAGCGC CCGGGCTATT GGCCGGGAGG TGGCTGGGTT CAAGGACCCG CGACTTGTC AGGACCCCGG AAGGGGAGG GGGGTGGGC AGCTCCACG  
CTAAATCGCG GGCCCGATAA CCGGCCCTCC ACCGACCAA GTTCTGGCC GTGAACAGT TCCTGGGGC TTCCCTCTCC CCCCACCCCG TCGGAGGTG  
301 TGCCAGCGGG GACTTGGGG AGTCCTTGGG GATGGCAAAA ACCTGACCTG TGAAGGGGAC ACAGTTTGG GGTGAGGGG AAGAAGGTTT GGGGTTCG  
ACGGTCGCC CTGAACCCCC TCAGGAACCC CTACCGTTT TGGACTGGAC ACTTCCCCTG TGTCAAACCC CCAACTCCCC TTCTTCCAAA CCCCCAAGC  
401 CTGTGCCAST GGAGAGGAAG CTGATAAGCT GATAACCTGG GCGCTGGAG CACCACCTAT CTGCCAGAG GNNNTGTA GCTGGGGTG GGGTGTGCAC  
GACACGGTSA CCTCTCCTC GACTATTGCA CTATTGGACC CCGACCTCG GTGGTGAATA GACGCTCTCC CNNNNACCAT CGACCCCCAC CCCACACGT  
501 ACGGCAGCAG GATTGAATGA AGGCCAGGGA GGCAGCACCT GAGTGCTTG ATGTTGGGG ACAGGAAGGA CGAGCTGGG CAGAGACGTG GGGATGAAGG  
TGCGGTGTC CTAACCTACT TCCGGTCCCT CCGTCGTGGA CTCACGAACG TACCAACCCC TGTCCTTCT GTCTCTGCAC CCCTACTTCC  
601 AAGCTGTCTT TCCACAGCCA CCTTCTCCC TCCCGCCCTG ACTCTCAGCC TGCTATCTC TTCTAGAAATG TCCTGCCTGG CTGTGGCTTC TCCTGTCCCT  
TTCGACAGGA AGGTGTCGGT GGAAGAGGG AGGGGGGAC TGAGAGTCCG ACCGATAGAG AAGATCTTAC AGGACGGACC GACACCGAAG AGGACAGGGA  
1 E C P A W L W L L S L  
^exon 2  
701 GCTGTGCTC CCTCTGGGC TCCAGTCTT TCCAGTCTT GGGGCCCCA CCACGCCCTA TCTGTACAG CCGAGTCTG GAGAGGTACC TCTTGGAGG CAAGGAGGCC  
CGACAGCGAG GGAGACCCCG AGGATCAGGA CCGCGGGGT GGTGGGAGT AGACACTGTC GGCTCAGGAC CTCCTCATGG AGAACCTCCG GTTCTCTCCG  
13 L S L P L G L P V L G A P P R L I C D S R V L E R Y L L E A K E A  
801 GAGAAATATCA CCGTGAGACC CCTTCCCCAG CACATTCCAC AGAATCAGC CTCAGGGCTT CAGGGAACCTC CTCCCAGATC CAGGAACCTG GCACCTTGGT  
CTCTTATAGT GCCACTCTGG GGAAGGGGTC GTGTAAGGTG TCTTGAGTGC GAGTCCCGAA GTCCCTTGAG GAGGTCTAG GTCCCTTGAC CGTGAACCAA  
46 E N I T

7

7

7

7

7

7

7

7

7

7

7

7

C  
A  
C  
T.  
D  
T.  
T.  
E  
E  
T  
U  
D  
O

# FIGURE 1C

2001 AGGAGTACAG GAACTGTCCG TATTCCTTCC CCTTCTGTGG CACTGCAGCG ACCTCCTGTT TTCTCCTTGG CACAAGGAAG CCATCTCCCC TCCAGATGCG  
 TCCTCATGTC CTTGACAGGC ATAGGAAGG GGAAGACACC GTGACGTGCG TGGAGGACAA AAGAGGAACC GTCTTCTTTC GGTAGAGGG AGGTCTACGC  
 K E A I S P P D A  
 ^exon 5

1

2101 GCCTCAGCTG CTCCACTCCG AACAACTACT GCTGACACTT TCCGCRAACT CTTCGAGTC TACTCCAATT TCCTCCGGG AAAGCTGAAG CTGTACACAG  
 CGGAGTCGAC GAGGTGAGGC TTGTTAGTGA CGACTGTGAA AGCGTTTGA GAAGGCTCAG ATGAGGTTAA AGGAGGCCCC TTTCGACTTC GACATGTGTC  
 10 A S A A P L R T I T A D T F R K L F R V Y S N F L R G K L K L Y T G

2201 GGGAGGCCCTG CAGGACAGGG GACAGATGAC CAGGTGTGTC CACCTGGGCA TATCCACCAC CTCCCTCACC AACATTGCTT GTGCCACACC CTCCCCCGCC  
 CCCTCCGGAC GTCCTGTCCC CTGTCTACTG GTCCACACAG GTGGACCCGT ATAGGTGGTG GAGGAGTGG TTGTAACGAA CACGGTGTGG GAGGGGGCGG  
 44 E A C R T G D R O

2301 ACTCCTGAAC CCCGTCGAGG AGCTCTCAG  
 TGAGGACTTG GGGCAGCTCC TCGAGAGTC

FIGURE 2

1 ATGGGGGTGC ACGAATGTCC TGCCTGGCTG TGGCTTCTCC TGTCCTCTGCT GTGCTCCCTCT CTGGGCTCTCC CAGTCTCTGG CGCCACCA CGCTCATCT  
TACCCCCACG TGCTTACAGG ACGGACCGAC ACCGAAGAGG ACAGGGACGA CAGCAGGGA GACCCGGAGG GTCAGAGCC GCGGGTGGT GCGGAGTAGA  
1 M G V H E C P A W L W L L L S L L S L P L G L P V L G A P P R L I C

101 GTGACAGCG AGTCCTGGAG AGGTACCTCT TGGAGGCCAA GGAGGCCGAG AATATCAGA CGGCTCTGC CGAACACTGC AGCTTGAATG AGAATATCAC  
CACTGTGGC TCAGGACCTC TCCATGGAGA ACCTCGGCTC TTATAGTGT GCGGACAGG GCTTGTGACG TCGAATTTAC TCTTATAGTG  
35 D S R V L E R Y L L E A K E A E N I T T G C A E H C S L N E N I T  
^\*

201 TGTCCACAG ACCAAAGTTA ATTTCTATGC CTGGAAGAGG ATGAGGTCA GGAGCAGGC CGTAGAGTC TGGCAGGGCC TGGCCTCTGCT CTCGGAAGCT  
ACAGGGTCTG TGGTTTCAAT TAAAGATACG GACCTTCTCC TACCTCCAGT CCGTCGTCG GCATCTTCAG ACCGTCCCG ACCGGGACGA GAGCCTTGA  
68 V P D T K V N F Y A W K R M E V R Q Q A V E V W Q G L A L L S E A

301 GTCCTGGGG GCCAGGCCCT GTTGGTCAAC TCTTCCACG CGTGGGAGCC CCTGCAGCTG CATGTGGATA AAGCCGTGAG TGGCCTTCCG AGCCTCACCA  
CAGGACGCC CCGTCCGGGA CAACCAAGTTG AGAAGGTGCG GCACCCCTCGG GGACGTCGAC GTACACCTAT TTCGGCAGTC ACCGGAAGCG TCGGAGTGGT  
101 V L R G Q A L L V N S S Q P W E P L Q L H V D K A V S G L R S L T T  
^\*

401 CTCTGCTTCG GGCTCTGGGA GCCCAGAAGG AAGCCATCTC CCTCCAGAT GCGGCTCTCAG CTGCTCCACT CCGAACAAATC ACTGCTGACA CTTTCCGCCAA  
GAGACGAAGC CCGAGACCTT CCGGTCTTCC TTCCGTAGAG GGGAGGTCTA CCGCCGAGTC GACGAGGTGA GGCTTGTAG TGACGACTGT GAAAGGCGTT  
135 L L R A L G A Q K E A I S P P D A A S A A P L R T I T A D T F R K  
^Polymorphism in CRL 1609 (A or C). A changes the protein sequence to K.

501 ACTCTTCCGA GTCTACTCCA ATTTCTCTCC GGGAAAGCTG AAGCTGTACA CAGGGGAGGC CTGCAGGACA GGGGACAGAT GACCA  
TGAGAAGGCT CAGATGAGGT TAAAGGAGGC CCGTTTCGAC TTCGACATGT GTCCCTCCG GACGTCCTGT CCCCTGTCTA CTGGT  
168 L F R V Y S N F L R G K L K L Y T G E A C R T G D R O



### FIGURE 3

human MGVHECPAWLWLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAE

\*\*\*\*\*

chepo MGVHECPAWLWLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAE

human NITTGCAEHCSLNENITVPDTKVNIFYAWKRMEVGQQAVEVWQGLALLSEA

\*\*\*\*\*

chepo NITTGCAEHCSLNENITVPDTKVNIFYAWKRMEVRQQAVEVWQGLALLSEA

human VLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLRLALGAQKEAISPPD

\*\*\*\*\*

chepo VLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLRLALGAXKEAISPPD

human AASAAPLRTITADTFRKLFRVYSNFLRGKCLKLYTGEACRTGDR

\*\*\*\*\*

chepo AASAAPLRTITADTFRKLFRVYSNFLRGKCLKLYTGEACRTGDR

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 4

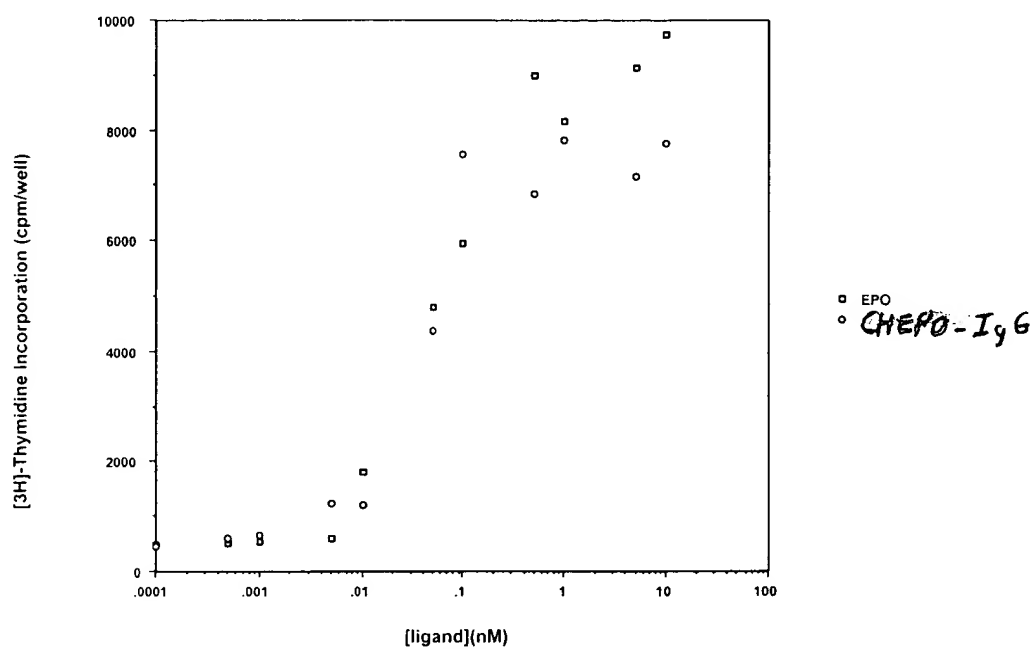


FIGURE 5

